

L'intelligence artificielle peut-elle devenir un outil convivial?

ou doit-on immédiatement arrêter toute recherche en IA?

Romain Couillet*, **Pierre-Olivier Amblard**[°], **Denis Trystram***, **Achille Baucher***

[°]GIPSA-lab

*LIG-lab, Université Grenoble-Alpes

romain.couillet@univ-grenoble-alpes.fr, pierre-olivier.amblard@gipsa-lab.grenoble-inp.fr,
denis.trystram@imag.fr, achille.baucher@uiv-grenoble-alpes.fr

Résumé

Depuis les premiers succès des réseaux de neurones profonds au début des années 2010, le secteur de l'intelligence artificielle (IA) et son bilan carbone croissent à une vitesse plus intense que l'ensemble du monde numérique, lui-même crédité d'une expansion annuelle de 6 à 9% (selon les chiffres). À l'heure où l'urgence climatique exige un repli annuel de la dépense carbone à hauteur de 5 à 7% afin de tenir les accords de Paris et d'éviter une hausse globale de la température terrestre de 2°C, la question se pose du bilan des gains et pertes pour la société induits par l'activité économique de l'IA.

Plus spécifiquement, notre interrogation consiste ici à nous demander si de la recherche en IA peuvent émerger des outils « conviviaux », au sens d'Illich, c'est-à-dire qui améliorent l'adaptation de l'humanité à son milieu sans créer de dépendance à la technologie et sans induire directement ou indirectement d'inégalités dans la population – susceptibles d'accroître les difficultés d'adaptation d'une partie de la population concernée. Un deuxième enjeu consiste alors à engager chercheurs et industries en IA, dont beaucoup admettent aujourd'hui se trouver en dissonance cognitive face à la perte de sens et au caractère écologiquement aggravant de leur recherche au quotidien, dans une transition souple vers le développement de ces outils conviviaux.

Mots-clés

Intelligence artificielle, société conviviale, recherche.

Déclaration des auteurs

L'ordre des auteurs est alphabétique et ne correspond donc pas à la part de contribution.

1. État des lieux de l'IA en 2021

1.1. 2012 ou l'année de la renaissance des vieux rêves cybernétiques

L'imaginaire d'un monde cybernétique, dans lequel non seulement l'énergie métabolique d'Homo Sapiens mais également une partie de sa mémoire, de son savoir et de ses prises de décision seraient délégués à des machines, remonte aux premiers transistors. Les limites de tels univers sous contrôle des machines ont été longtemps anticipées à divers degrés d'analyse littéraire, philosophique [1] ou fictionnel [2]. Néanmoins, de l'après-guerre à 2012, les faibles progrès de l'IA ont éloigné la perspective d'un monde de robots humanoïdes et ont vraisemblablement étioilé notre vigilance face aux alertes lancées dans les années 70.

La donne change en effet en 2012 avec le premier exploit réalisé, au prix de calculs massivement parallélisés, par un réseau de neurones artificiel dans une compétition de reconnaissance d'images [3]. L'emballement scientifique mondial, en à peine dix ans, donne soudain naissance à des algorithmes aux capacités sur-humaines en traitement des images et du langage naturel. Dans un contexte écologique par ailleurs toujours plus pressant en toile de fond, un nouvel imaginaire se crée de l'avènement possible d'une IA réparatrice des maux de la planète [4]. De nombreux champs disciplinaires s'emparent alors de l'outil qui leur promet un élan nouveau, et ce en dépit d'un manque central de garanties théoriques, conséquence du fonctionnement en « boîte noire » des algorithmes d'apprentissage profond.

1.2. Les usages de l'IA en 2021

Peu de travaux font un état des lieux concret de la distribution exacte des usages de l'IA. Les recoupements d'études nous permet néanmoins d'identifier les systèmes de filtrage de messages (emails, réseaux sociaux), de recommandations publicitaires ainsi que les usages commerciaux ou à vue de télésurveillance comme les plus gros champs d'applications, en volume, de l'IA. S'il existe bien des usages environnementaux (suivi et recensement de la biodiversité, traitement en masse de données d'évolutions écologiques), ces outils n'émergent en général que sous la forme de sous-produits secondaires aux investissements à visée productive industrielle. Au bilan, au même titre que la robotique ou les outils de traitement du signal à visée médicale bénéficient indirectement des avancées impulsées par des exigences militaires ou commerciales, l'opportunité de nouvelles applications en IA à impact environnemental positif surcompense-t-elle les coûts environnements attenants au développement de l'outillage industriel de l'IA?

Aujourd'hui, le taux de croissance annuel de l'empreinte carbone de l'ensemble du domaine numérique est de 6 % à 9 % [5]. Par ailleurs, la quantité d'opérations nécessaires à l'entraînement d'un réseau de neurones profonds, au cœur de la révolution de l'IA depuis 2012, double tous les 3,4 mois (soit une croissance de 200 % annuelle du nombre de « petaflops » requis) [6]. Les réseaux de neurones les plus avancés en 2020 coûtaient déjà 1,300,000 kWh à entraîner (soit environ 130,000€ au prix de l'électricité en France), ou l'équivalent de 500 TeqCO₂ en terme de bilan carbone (soit le cycle de vie de 3 SUVs) [7]. Ces chiffres n'ont de cesse d'augmenter au point d'atteindre aujourd'hui une saturation: pour diviser par deux le taux d'erreurs d'apprentissage, on observe empiriquement qu'il faut multiplier par 2⁹ (environ 500) le coût de l'entraînement [8]. Si certains réseaux de neurones sont aujourd'hui capables de certaines prouesses leur permettant de rivaliser avec le cerveau humain, il est important d'insister sur le fait qu'il ne s'agit essentiellement que d'applications aux contours limités (vision, certaines opérations de langage, apprentissage par renforcement pour les jeux) et requérant une quantité d'énergie incommensurable avec l'énergie métabolique nécessaire aux neurones biologiques humains (le cerveau consomme 20 % des 2 à 5 kWh que nous ingérons quotidiennement). Un premier mur théorique semble ainsi se dresser devant l'évolution future d'une IA à l'une empreinte carbone en croissance vive, aux usages et champs d'actions restreints, et dont le bénéfice environnemental net est fortement négatif.

Une autre promesse de l'IA s'inscrit dans le développement futur d'un triptyque 5G-IoT-IA qui permettrait de reporter, transmettre et traiter à la volée un grand nombre de données attendant aux secteurs de la santé, de la domotique, de l'automatisation des villes, de l'agriculture ou de l'évolution des territoires, ceci permettant dans l'idéal d'optimiser simultanément la consommation des structures et le bien-être collectif. Aucune étude ne reporte cependant de bénéfice net à ce déploiement d'instruments connectés, dont il faudra amortir les coûts de production, d'installation et de consommation. Ce dernier permet cependant de soulever

les questions de la pertinence sociétale et de notre capacité de résilience consécutive à l'installation de tels dispositifs high-tech, précisément à l'heure de préparer une ère post-thermo-industrielle dont la consommation énergétique ne pourra que fortement s'infléchir.

2. Recherche scientifique, IA et convivialité

2.1. Le contrôle du savoir

Illich anticipe en 1973 [1] l'effondrement de la société thermo-industrielle et dépeint les conditions nécessaires à l'instauration de (ou au retour à) une société post-industrielle résiliente, à savoir capable de supporter la décroissance marquée des sources d'énergie thermique. Illich confine cette résilience à la notion de « convivialité » des « outils » (biens et services) de la civilisation. Est dit convivial tout outil qui augmente le contrôle de l'humain sur son environnement sans pour autant créer de dépendance ni d'accroissement des inégalités sociales (qui ont tous deux pour résultat une diminution de la résilience). Illich pose en exemple la bicyclette comme outil de transport maximalement convivial (elle étend le rayon d'action, est accessible et aisément réparable par tous, peut être temporairement remplacée par la marche, etc.). A contrario, une technologie bénéfique pour une communauté mais qui impacte négativement l'espace de vie et des possibles d'une autre (telle l'automobile) n'est pas conviviale.

Pour Illich, la production et le partage du savoir eux-mêmes (la recherche en particulier) ne sont pas conviviaux. La recherche s'approprie le terrain du savoir-faire collectif pour le segmenter en spécialités indépendantes sous le « monopole radical » d'un petit groupe d'élites. Ce monopole radical exclue (parfois par voies légales comme dans le cadre de la médecine) l'instruction autodidacte et délègue l'exploitation et la définition du savoir à des groupes distincts qui s'éloignent tous, par effet de spécialisation, d'une vision large de l'environnement devenue inaccessible à tous. Illich illustre ce propos par les investissements massifs en chirurgie de pointe, réservée à une minorité, au détriment d'un apprentissage partagé des notions pratiques d'hygiène et de soins de base au coût social généralement inférieur. Un exemple marquant est celui de la consommation depuis l'ère industrielle d'aliments à indice glycémique élevé qu'on sait aujourd'hui être une cause majeure de nombreuses maladies dites « de civilisation » [9] au coût économique et social (de « pleine santé ») importants, et qu'il est pourtant théoriquement simple de bannir par une instruction collective élémentaire.

À la même époque qu'Illich, le mathématicien Alexandre Grothendieck, alors récemment médaillé Fields, décide d'abandonner toute recherche scientifique, invoquant sa vacuité et ses dangers [10]. Au compte de ces derniers, Grothendieck rejoint Illich sur l'enchaînement du chercheur à son « outil », l'utilisation de la recherche à des fins environnementalement mortifères ou à des fins militaires inéthiques. Ce savoir placé entre les mains d'une élite au travail parfois ressenti comme stérile est non convivial en cela qu'il isole le savoir d'une population dès lors plus en mesure d'exprimer sa curiosité créatrice. Steigler, qui nous est plus contemporain, va même plus loin en avançant que les médias grand public, dont la survie repose sur l'efficacité du martellement publicitaire, dépossède totalement la population de sa capacité même de création et d'esprit critique [11]. Indirectement, ces auteurs mettent en lumière l'incapacité de résilience d'une population occidentale dépendante des machines alimentées par les énergies thermiques, démunie de sa puissance d'agir (ce que Spinoza identifie au passage comme la source du bien-être [12]), en particulier collectivement, à l'orée d'une ère nouvelle où l'énergie métabolique (entraide, artisanat, agriculture) devra compenser en grande partie l'épuisement des énergies fossiles.

2.2. Faut-il dès lors abandonner toute recherche en IA ?

On oppose naturellement à Illich et Grothendieck que la recherche a permis à nombre d'entre nous de survivre à des maladies autrefois incurables et augmenté le confort global moyen (ce second argument étant mis en doute par les philosophes du bonheur). Se dessine ici en réalité une ligne rouge, qu'Illich nomme convivialité, qu'un outil ne saurait dépasser au risque de devenir socialement contre-productif. Cette ligne rouge fluctue aussi dans le temps et l'espace: un outil peut être socialement acceptable dans un contexte de

L'intelligence artificielle peut-elle devenir un outil convivial ?

ressources abondantes et d'impact environnemental restreint, mais être non-convivial ailleurs et en d'autres temps. L'urgence écologique de ce début de 21^e siècle exige de mettre rapidement en œuvre des mesures de forte contraction économique (précisément de consommation énergétique fossile, ce qui en ordre de grandeur est essentiellement équivalent) [13] et réduit progressivement le périmètre décrit par la ligne rouge de convivialité, notamment de la recherche en IA.

Si on confronte l'état des avancées scientifiques et des usages de l'IA aux contraintes environnementales et à la nécessité d'une résurgence de résilience collective, il paraît difficile de justifier d'une nécessité de continuer à développer un savoir expert, particulièrement en réseaux de neurones profonds ou en matériel (parallélisation massive de cœurs GPU ou TPU dans des serveurs de calcul très consommateurs). Il ne s'agit néanmoins pas de faire table rase du long cumul de progrès techniques, et particulièrement des réseaux de neurones déjà entraînés aujourd'hui et capables de fonctionner sur une architecture matérielle légère, ni de la palette d'outils conviviaux et de méthodes de raisonnements que la recherche en informatique et mathématiques appliquées nous ont apportée jusque là. Rappelons notamment que dans le domaine des technologies de l'information et des communications (TIC), des théories puissantes telles que la théorie de l'information de Shannon ou le traitement statistique du signal ont révolutionné la science du stockage, de l'échange et du traitement de l'information (mathématique mais aussi sociale) et nous permettent aujourd'hui d'opérer ces technologies au plus près de leur optimalité énergétique théorique.

La recherche scientifique en IA pourrait alors engager un virage qui l'éloignerait de la quête de performances « quoi qu'il en coûte » pour se poser la question de ses usages dans des contextes énergétiquement contraints et permettant d'augmenter la métrique de convivialité (et ainsi de résilience) de l'outil; en somme, préparer la mutation de l'IA vers l'ère post-industrielle. Cette nouvelle vision de l'IA pose d'emblée la question de la « bonne métrique » qu'elle tendrait dès lors à maximiser. De telles réflexions sont actuellement menées dans le cadre des technologies dites « low-techs » qui, bien plus que des outils basés sur des techniques oubliées ou confinées aux pays en développement, s'ancrent aujourd'hui précisément dans une recherche de convivialité et de résilience maximale, implémentée localement mais partagée largement, et ce bien sûr en anticipation des contraintes écologiques et sociétales à venir [14].

3. Les pistes d'une IA conviviale

3.1. De l'IA high-tech à une économie low-tech

Le mouvement low-techs prend en effet une envergure telle qu'il s'imisce aujourd'hui dans certaines formations universitaires en parallèle des cours plus conventionnels sur l'économie verte. Les low-techs visent la préparation de sociétés bas carbone dans leur usage minimaliste de ressources énergétiques et minérales – ou du moins dans un fonctionnement économique circulaire – et *s'inspirent* de pratiques historiques révolues ou mises en place dans les pays du sud. Si le but est d'établir une position de résilience sur le long court qui exige de fait une capacité de régénération de sociétés entièrement low-techs, l'urgence climatique et les bases de connaissances modernes ne permettent pas d'exclure la possibilité de développer des low-techs au moyen d'outils high-techs. Ici encore se dresse cependant une ligne rouge délicate à définir et à ne pas franchir, et qui sert d'alibi à l'économie de croissance verte.

Des exemples concrets d'IA moderne au service des low-techs et de la résilience existent néanmoins. Le plus intéressant à nos yeux est le développement par l'équipe du docteur d'Acremont [15] d'un arbre de décisions manuscrit permettant le diagnostic amélioré du bilan médical d'enfants en Tanzanie en l'absence de professionnels de santé et d'électricité. L'arbre a été développé suite à l'élaboration d'un algorithme d'apprentissage sur la base de données collectées localement. L'enjeu du concours des méthodes high-tech pour la conception de l'outil low-tech aura été ici d'analyser finement les réponses d'une gamme d'algorithmes et de retenir celui dont l'explicabilité était la plus simple. On retrouve ainsi un axe majeur des métriques de convivialité qui tend à maximiser, et même à augmenter, la maîtrise de l'utilisateur (le service de soin local) sur son outil (l'arbre de décision). L'approche par réseaux de neurones ne se révèle au passage pas plus performante et n'est surtout pas suffisamment explicable pour être acceptable.

3.2. L'intelligence artificielle au service de l'intelligence collective ?

L'exemple précédent ouvre la voie à une nouvelle forme de recherche, ou du moins d'usage, de l'IA qui exploite la machine à des fins descriptives de comportements de systèmes que l'humain peut alors s'approprier. Dans ce cadre, l'apprentissage profond – dont on espère peu des progrès d'interprétation dans les années à venir – sortirait de la boucle au profit des arbres de décisions et des méthodes simples dérivées de régressions linéaires bien comprises (SVM, régression logistique, réseaux echo-state en dynamique, etc.). Ses outils ont l'avantage de fonctionner sur la base de peu de données, sur des laptops conventionnels, d'être théoriquement bien compris et accessibles à une majorité de non-initiés. Il s'agira toutefois d'apprendre à faire mieux parler ces algorithmes et de viser une finalité résiliente et auto-régénératrice permettant de se passer à terme de l'outil informatique.

Pourquoi pas?

Bibliographie

- [1] Illich, I., Giard, L., & Bardet, V. (1973). *La convivialité*. Paris: Editions du Seuil.
- [2] Asimov, I., & Rosenthal, J. (1984). *Fondation*. [successivement] Denoël: Presses de la Cité.
- [3] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 1097-1105.
- [4] <https://www.weforum.org/agenda/2018/01/8-ways-ai-can-help-save-the-planet/> (voir aussi Wu, J., Guo, S., Li, J., & Zeng, D. (2016). Big data meet green challenges: Big data toward green applications. *IEEE Systems Journal*, 10(3), 888-900.)
- [5] The Shift Project (2019). Lean ICT – Towards Digital Sobriety. Report of the working group directed by Hugues Ferreboeuf for the think tank the Shift Project.
- [6] <https://openai.com/blog/ai-and-compute/>
- [7] Patterson, D., Gonzalez, J., Le, Q., Liang, C., Munguia, L. M., Rothchild, D., ... & Dean, J. (2021). Carbon emissions and large neural network training. *arXiv preprint arXiv:2104.10350*.
- [8] Thompson, N. C., Greenewald, K., Lee, K., & Manso, G. F. (2020). The computational limits of deep learning. *arXiv preprint arXiv:2007.05558*.
- [9] Taubes, G. (2007). *Good calories, bad calories*. Anchor.
- [10] Grothendieck, A. (2016). Allons-nous continuer la recherche scientifique?. *Ecologie politique*, (1), 159-169.
- [11] Stiegler, B. (2010). For a new critique of political economy. *Polity*.
- [12] Lenoir, F. (2017). *Le miracle Spinoza: une philosophie pour éclairer notre vie*. Fayard.
- [13] The Shift Project (2017). Modeling global emissions trajectories compatible with a +2°C carbon budget.
- [14] Bihouix, P. (2014). *L'Âge des low tech. Vers une civilisation techniquement soutenable*. Média Diffusion.
- [15] D'Acremont V. (2021). Technologies et santé : Quels compromis entre éthique, environnement et climat ? Analyse réflexive et expérience de terrain. <https://www.youtube.com/watch?v=ni-I AxepSs>